



测绘学报

Acta Geodaetica et Cartographica Sinica

ISSN 1001-1595, CN 11-2089/P

《测绘学报》网络首发论文

题目: Causal-GCNM: 顾及缺失值的因果图时空预测网络
作者: 王培晓, 张彤, 聂士超, 杨瑾萱, 王天娇
网络首发日期: 2022-09-15
引用格式: 王培晓, 张彤, 聂士超, 杨瑾萱, 王天娇. Causal-GCNM: 顾及缺失值的因果图时空预测网络[J/OL]. 测绘学报.
<https://kns.cnki.net/kcms/detail/11.2089.p.20220913.1951.010.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

Causal-GCNM：顾及缺失值的因果图时空预测网络

王培晓¹，张彤¹，聂士超²，杨瑾萱³，王天娇¹

1. 武汉大学测绘遥感信息工程国家重点实验室，湖北 武汉 430079；2. 武汉大学资源与环境科学学院，湖北 武汉 430079；3. 武汉大学遥感信息工程学院，湖北 武汉 430079

A Causal Graph Convolutional Network Considering Missing Values for Spatiotemporal Prediction

WANG Peixiao¹, ZHANG Tong¹, NIE Shichao², YANG Jinxuan³, WANG Tianjiao¹

1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; 2. School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China; 3. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China;

Abstract: Spatiotemporal prediction is one of the basic research topics of geographic spatiotemporal big data mining. There are many attempts to predict spatiotemporal state of unknown systems using various deep learning algorithms. However, most existing prediction models are only tested on spatiotemporal data assuming no missing data entries, ignoring the impact of missing values on the prediction results. In the actual scenarios, data missing is an inevitable problem due to sensor or network transmission failures. Therefore, we propose a novel Causal Graph Convolutional Network considering Missing Values (Causal-GCNM) for spatiotemporal prediction. The proposed model can automatically capture missing patterns in the spatiotemporal data, enabling the Causal-GCNM model to directly complete the spatiotemporal prediction task without additional interpolation. The proposed model was validated on three real spatiotemporal datasets (traffic flow dataset, PM_{2.5} monitoring dataset, and temperature monitoring dataset). Experimental results show that the Causal-GCNM model has good prediction performance under four missing scenarios (20% random missing, 20% block missing, 40% random missing, 40% block missing), and outperforms ten existing baseline methods in terms of prediction accuracy and computational efficiency.

Keywords: Geographic spatiotemporal big data mining; causal convolution network; graph convolution network; spatiotemporal prediction; spatiotemporal data missing

Foundation support: National Key R&D Program of China (International Scientific & Technological Cooperation Program) under grant 2019YFE0106500; National Natural Science Foundation of China under grant 41871308; the Fundamental Research Funds for the Central Universities

摘要：时空预测是地理时空大数据挖掘的基础研究命题。目前，多种模型用于预测未知系统的时空状态。然而，存在的大多数预测模型仅没有缺失数据的时空数据集上进行测试，忽略了缺失值对预测结果的影响。在真实场景中，由于传感器或网络传输故障，数据缺失是一个不容忽视的问题。鉴于此，本文提出了一种新颖的顾及缺失值的因果图卷积网络（Causal Graph Convolutional Network considering Missing Values, Causal-GCNM）用于时空预测。Causal-GCNM 模型可以自动的捕捉时空数据中的缺失模式，使得 Causal-GCNM 模型在不需要借助额外插值算法的前提下，可以直接完成时空预测任务。本文提出的模型在三种真实的时空数据集（交通流数据集、PM_{2.5}监测数据集、气温监测数据集）得到了验证。实验结果表明，Causal-GCNM 模型在四种缺失条件（20%随机缺失、20%块状缺失、40%随机缺失、40%块状缺失）下仍然具有较好的预测性能，并在预测精度和计算效率两类指标上优于 10 种存在的基线方法。

关键词：地理时空大数据挖掘；因果卷积网络；图卷积网络；时空预测；时空数据缺失

中图分类号：P208

文献标识码：A

基金项目：国家重点研发计划“政府间国际科技创新合作/港澳台科技创新合作”重点专项(2019YFE0106500)；国家自然科学基金项目(41871308)；中央高校基本科研业务费专项资金资助

时空预测是地理时空大数据挖掘的基础研究命题,在时间和空间上对未知的系统状态进行预测^[1-4]。目前,时空预测已被广泛应用于天气预报^[5-6]、智能交通^[7-8]、地震预警^[9]等实际应用中,从而辅助有关部门及时开展应急决策,防止和减轻各种灾害带来的社会经济损失。时空预测技术的性能严重依赖于时空数据。幸运的是,随着物联网的快速发展,许多传感器数据呈现爆发式增长,为时空预测相关的研究提供了重要的数据源^[10-12]。

时空预测的关键在于捕捉时空数据中的复杂时空相关性。目前,国内外相关学者采用不同方法展开了研究以提高时空预测的性能,例如,时空 K 临近模型 (Spatiotemporal k-nearest neighbor, ST-KNN)^[13-14]、时空残差网络 (Spatiotemporal residual networks, ST-ResNet)^[15]、卷积门控单元网络 (Convolutional Gate Recurrent Unit, ConvGRU)^[16]等。但 ST-KNN、ST-ResNet 和 ConvGRU 模型主要针对欧式数据结构设计,难以应用到非欧数据结构中。近年来,考虑到研究对象可能存在非欧数据结构,图卷积神经网络 (Graph convolutional neural network, GCN)^[17]逐渐应用到时空数据建模中并取得了最优的预测性能,例如时间图卷积网络 (Temporal graph convolutional network, T-GCN)^[18]、时空图卷积网络 (Spatiotemporal graph convolutional network, ST-GCN)^[19]和残差图卷积长短期记忆 (Residual graph convolution long short-term memory, RGC-LSTM) 网络^[20]等。尽管国内外相关学者针对时空预测做了大量的工作,但仍存在不足。具体而言, T-GCN、ST-GCN 和 RGC-LSTM 等预测模型仅在没有缺失数据的情况下对时空数据进行测试,即为非缺失数据预测建模^[21-22]。然而,在实际环境中,由于采集技术的各种问题和故障,数据缺失普遍存在并严重影响时空预测模型的性能^[23-25]。存在的预测模型主要采用两种策略处理缺失的时空数据^[26]: (1)构建预测模型之前采用时空插值模型填补缺失的数据; (2)删除含有缺失数据的时间序列。前者增加了额外的计算负担并且填补精度直接影响到模型的预测性能^[8,27-28],后者可能导致模型的训练数据不足且无法获取可靠的时空模式^[20,29-30]。

除了基于非缺失数据建模以外,部分学者顾及时空数据的缺失模式直接基于缺失数据建立时空预测模型,例如,矩阵分解和张量分解模型^[31-32]、谱图马尔可夫网络 (spectral graph Markov network, SGMN)^[33]、基于延迟机制的门控循环网络 (Gate recurrent unit with decay mechanism, GRU-D)^[34]、基于缺失数据的长短时记忆网络 (Long short-term memory network with missing data, LSTM-M)^[35]等。然而, SGMN、GRU-D 和 LSTM-M 等顾及缺失值的预测模型多是一种迭代类型的预测模型,迭代型的预测模型计算复杂度较高,不利于实时性要求高的预测应用。

与迭代类型的预测模型相比,因果卷积网络 (Causal convolution network, Causal-CN) 是一种新颖的非迭代类型的预测模型^[36]。已有研究表明 Causal-CN 具备运算速度快,精度高的优点^[37]。目前, Causal-CN 模型已被广泛应用在信号分类^[38]、序列预测^[39]等诸多领域,并获得了较高的预测性能。因此,本文尝试将 Causal-CN 模型应用到含有缺失值的时空预测任务中。针对本文的研究目标,传统的 Causal-CN 模型虽然具备上述诸多优点,但仍存在两点不足。首先, Causal-CN 模型是一种单纯的时间序列模型,难以挖掘空间对象之间的空间相关、空间依赖关系;其次, Causal-CN 模型并没有顾虑时空数据集中的缺失现象,无法识别时空数据集中常见的随机缺失及块状连续缺失模式。

鉴于此,本文提出了一种新颖的顾及缺失值的因果图卷积网络 (Causal Graph Convolutional Network considering Missing Values, Causal-GCNM) 模型。具体而言,本文的主要贡献总结如下:

- (1) 本文在 Causal-CN 模型的基础上扩展了图卷积操作,从而提出了一种新颖的端到端的时空预测模型 Causal-GCNM。Causal-GCNM 模型不仅可以应用在单纯的时间序列数据,还可以服务于地学领域的多种时空预测任务。
- (2) 在 Causal-GCNM 模型中,本文基于缺失时间窗口距离观测时间窗口的时间间隔,将缺失数据的缺失模式转换为缺失模式概率,从而处理不同缺失模式下的缺失值。缺失数据处理层使得 Causal-GCNM 模型在不需要借助额外插值算法的前提下,可以直接完成时空预测任务。
- (3) 本文采用三种真实的时空数据集 (交通流数据集、PM_{2.5} 监测数据集、气温监测数据集) 评估了 Causal-GCNM 模型的预测性能。此外,作者将采用的数据集和源码开源,以支持地学领域其他研究人员开展顾及缺失值的时空预测研究。

1 问题与定义

定义 1: 图 $G = \langle V, E, A \rangle$ 表示由监测站点抽象为的图结构, 其中 $V = \{v_i\}_{i=1}^n$ 表示图 G 中的 n 个节点, 即 n 个监测站点; E 表示节点之间的关联关系, 为简单起见, 节点之间的关联关系可用关系矩阵 $A \in \mathcal{R}^{n \times n}$ 表示。在不同场景中, 矩阵 A 可具备不同的含义, 例如, 在交通场景中, A_{ij} 可以表示节点 v_i 和节点 v_j 之间的拓扑连接关系; 在空气质量监测场景中, A_{ij} 又可以表示节点 v_i 和节点 v_j 之间的空间距离。

定义 2: 时空数据 x_i^t 表示节点 v_i 在时间窗口 τ_t 内监测到的数值, 例如单位时间内的交通流量、空气质量等。如图 1(a) 所示, 图 G 中的所有时空数据可表示为一个含有缺失数据的时空状态矩阵 $X \in \mathcal{R}^{n \times g}$, 其中 $x_i = \{x_i^t\}_{t=1}^g \in \mathcal{R}^{g \times 1}$ 表示节点 v_i 在所有时间窗口下的的时间序列, $x^t = \{x_i^t\}_{i=1}^n \in \mathcal{R}^{n \times 1}$ 表示所有节点在时间窗口 τ_t 内的空间序列。

定义 3: 0-1 矩阵 $M \in \mathcal{R}^{n \times g}$ 是一个仅包含元素 0 和 1 的矩阵, 用于区分时空状态矩阵 X 中的缺失值及观测值。若 $m_i^t = M_{it} = 0$, 则表示节点 v_i 在时间窗口 τ_t 内的监测值缺失, 即 x_i^t 为空。同理, $m_i \in \mathcal{R}^{g \times 1}$ 用于区分时间序列 x_i 中的缺失值及观测值; $m^t \in \mathcal{R}^{n \times 1}$ 用于区分空间序列 x^t 中的缺失值及观测值。

定义 4: 时间图序列 $TGS = \{G^t\}_{t=1}^g$ 表示图信息随时间不断改变的动态图序列, 本文特指图中节点的监测值随时间不断改变, 即 $G^t = \langle V, E, A, x^t, m^t \rangle$ 。如图 1c 所示, G^t 表示时间窗口 τ_t 内的图信息, 其中 x^t 和 m^t 的含义与定义 2 和定义 3 中的含义相同。

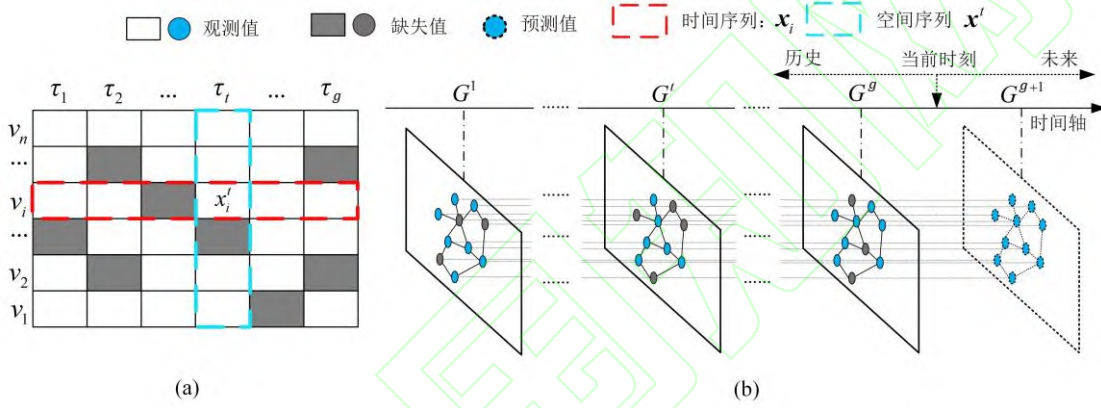


图 1 问题定义

Fig. 1 Problem definitions

本文的研究目标旨在建立一个函数模型 $\mathcal{F}(\cdot)$, 该模型可以从含有缺失值的时间图序列 TGS 中挖掘时空数据的模式, 从而准确的预测将来的时空数据。给定一个时间图序列 $TGS = \{G^t\}_{t=1}^g$, 其建模过程为

$$\hat{x}^{g+1} = \mathcal{F}(\{G^t\}_{t=1}^g; W) \quad (1)$$

式中 $\{G^t\}_{t=1}^g$ 表示含有缺失值的历史时间图序列; \hat{x}^{g+1} 表示将来 (预测) 的时空数据; $\mathcal{F}(\cdot)$ 表示本文提出的预测模型, 即 Causal-GCNM 模型; W 表示模型中可学习的参数。

2 顾及缺失值的因果图网络模型

本小节描述了 Causal-GCNM 模型的整体流程, 如图 2 所示, Causal-GCNM 模型主要由两个核心组件组成: 缺失数据处理层和时空依赖模式挖掘层。首先, 缺失数据处理层用于自动的挖掘时间图序列中的缺失模式 (随机缺失模式和连续块状缺失模式), 并处理不同缺失模式下的缺失值; 然后, 时空依赖模式挖掘层被用于刻画时空图序列中的时空相关性, 具体而言, 图卷积用于挖掘时空图序列中的空间相关性, 因果卷积用于挖掘时空图序列中的时间相关性; 最后, 通过模型输出 (预测值) 和观测真值的损失函数优化 Causal-GCNM 模型中的参数。

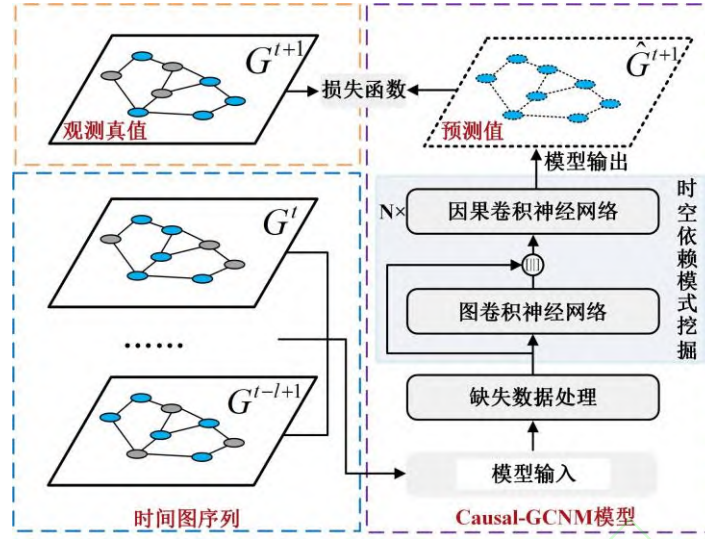


图 2 Causal-GCNM 模型的整体示意图

Fig. 2 Overall schematic of the Causal-GCNM model.

2.1 缺失数据处理层

缺失数据处理层的作用主要有两点，其一是识别缺失数据的缺失模式，其二是依据缺失模式处理数据中的缺失值。为了使得 Causal-GCNM 模型具备识别数据缺失模式的能力，我们引入了一个辅助量 $\mathbf{c}^t = \{c_i^t\}_{i=1}^n$ ，其计算方法为

$$c_i^t = \begin{cases} \tau_t - \tau_{t-1} + c_{i-1}^t & t > 1, m_i^{t-1} = 0 \\ \tau_t - \tau_{t-1} & t > 1, m_i^{t-1} = 1 \\ 0 & t = 1 \end{cases} \quad (2)$$

式中 c_i^t 表示节点 v_i 在时间窗口 τ_t 内的缺失值距离最近观测值的时间步长；当 $c_i^t = 1$ 且 $m_i^t = 0$ 时，节点 v_i 在时间窗口 τ_t 内的缺失模式趋向于是随机缺失；当 $c_i^t > 1$ 且 $m_i^t = 0$ 时，节点 v_i 在时间窗口 τ_t 内的缺失模式趋向于连续块状缺失。

在 \mathbf{c}^t 的基础上，本文进一步定义了 Causal-GCNM 模型处理缺失值的过程。其设计的思路来源主要有两点：(1) 在随机缺失模式下，缺失值与时间维度观测值的相关性要大于与空间维度观测值的相关性；(2) 当缺失模式趋于连续块状缺失时，缺失值与时间维度观测值的相关性将会逐渐下降，而预测值与空间维度观测值的相关性将会逐渐上升。基于这个思路，缺失值的估计值应该是空间维度观测值和时间维度观测值加权的结果，其权重与缺失模式相关。具体而言，Causal-GCNM 模型处理缺失值的计算方法为

$$\begin{cases} \mathbf{x}^t = \mathbf{m}^t \odot \mathbf{x}^t + (1 - \mathbf{m}^t) \odot (\delta^t \odot \mathbf{x}^{t:tm} + (1 - \delta^t) \odot \mathbf{x}^{t:sm}) \\ \delta^t = \exp\{-\max(0, \mathbf{W}_\delta \mathbf{c}^t + \mathbf{b}_\delta)\} \end{cases} \quad (3)$$

式中 $\mathbf{x}^{t:tm} \in \mathcal{R}^{n \times 1}$ 表示所有节点在时间维度上的最近观测值； $\mathbf{x}^{t:sm} \in \mathcal{R}^{n \times 1}$ 表示所有节点在空间维度上的最相似观测值； δ^t 表示由 \mathbf{c}^t 计算的缺失模式概率，即权重；当 δ^t 趋近 1，缺失数据的模式趋近随机缺失，当 δ^t 趋近 0，缺失数据的模式趋近连续块状缺失； \exp 表示指数函数； \max 表示最大值函数； $\mathbf{W}_\delta \in \mathcal{R}^{1 \times 1}$ 分别表示缺失数据处理组件中可以优化的参数。在缺失数据处理层，Causal-GCNM 模型虽然获得了空间最近观测值 $\mathbf{x}^{t:sm}$ 及时间最近观测值 $\mathbf{x}^{t:tm}$ ，但 Causal-GCNM 模型并不是依靠观测值直接填补时间窗口内的缺失值，而是依靠学习到的权重 \mathbf{W}_δ 自主的确定缺失值的处理策略。具体而言，Causal-GCNM 模型并不是真正的补全数据中的缺失值，而是依据最终的时空预测任务，基于数据的缺失模式自主的确定空间最近观测值 $\mathbf{x}^{t:sm}$ 及时间最近观测值 $\mathbf{x}^{t:tm}$ 的权重。

如图 3 所示，本文以节点 v_4 在时间窗口 τ_4 内的缺失值为例，进一步描述了缺失数据处理的流程。在时间维度上，我们直接搜索时间轴上的最近观测值，即 $\mathbf{x}_4^{4:tm}$ 。在空间维度上，我们首先计算了节点 v_4 与其他节点之间的空间相似性向量 $\mathbf{s}_4 = \{s_{4j}\}_{j=1}^n$ ，其中 s_{4j} 表示时间序列 \mathbf{x}_4 和时间序列 \mathbf{x}_j 的空间相似性，即 $s_{4j} = \|\mathbf{x}_4 - \mathbf{x}_j\|$ ；然后将 \mathbf{s}_4 中的元素递减排序，从而找到空间维度上的最近观测值，即 $\mathbf{x}_4^{4:sm}$ 。获得 $\mathbf{x}_4^{4:sm}$ 和 $\mathbf{x}_4^{4:tm}$

后，即可通过公式（3）求得节点 v_4 在时间窗口 τ_4 内的估计值。

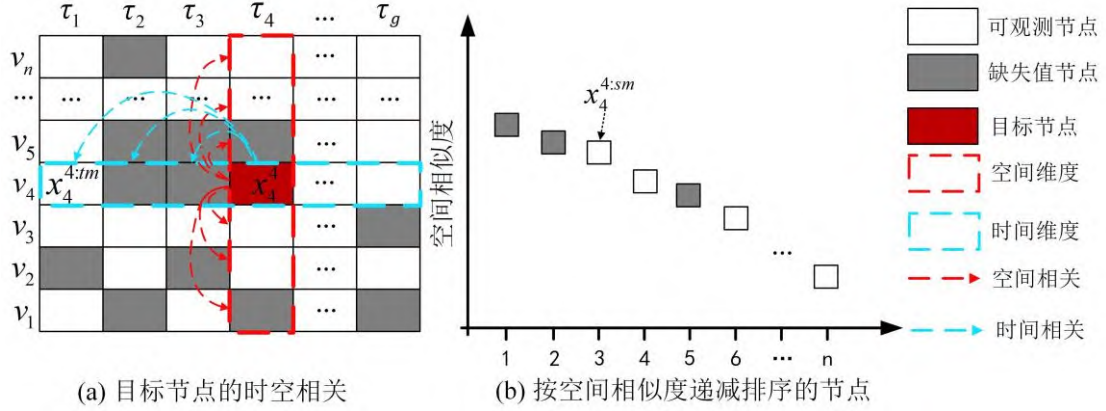


图 3 缺失数据处理组件示意图

Fig. 3 Illustration of missing data processing components.

2.2 时空依赖模式挖掘层

Causal-GCNM 模型的时空依赖模式挖掘层由单个图卷积神经网络和 N 个因果卷积神经网络组成。因此，本文首先简要介绍图卷积操作和因果卷积操作的基本原理（图 4）。在图卷积操作中，目标节点在空间维度上接受邻域节点的状态信息。在因果卷积中，目标节点在时间维度上接受自身的历史状态信息。通过整合因果卷积和图卷积操作即可捕捉时间图序列的时空模式[21]。考虑到传统的因果卷积操作难以捕捉长序列，本文进一步采用因果膨胀卷积操作替换了传统的因果卷积操作。如图 4(b)、图 4(c)所示，针对一个长度为 8 的时间图序列，卷积核等于 2 时，往往需要 7 层的传统因果卷积操作才可捕捉整条序列的信息。对于因果膨胀卷积操作而言，当卷积核和膨胀因子等于 2 时，仅需要 3 层的因果膨胀卷积操作即可捕捉整条序列的信息。图卷积、因果卷积、因果膨胀卷积的计算方法分别为

$$\star^g v_i = \sum_{j \in N_i} A_{ij} x_j W^g \quad (4)$$

$$(\star^c x_i)(t) = \text{sum} \left(W^c \odot (x_i^t, x_i^{t-1}, \dots, x_i^{t-k+1}) \right) \quad (5)$$

$$(\star_d^c x_i)(t) = \text{sum} \left(W_d^c \odot (x_i^t, x_i^{t-d}, \dots, x_i^{t-(k-1)d}) \right) \quad (6)$$

式中 \star^g 表示图卷积操作； \star^c 表示因果卷积操作； \star_d^c 表示膨胀因子为 d 的因果膨胀卷积操作； \odot 表示 Hadamard 积，即逐元素相乘； A_{ij} 图中节点 v_i 和 v_j 的连接关系； N_i 表示节点 v_i 的空间邻域集合； $(x_i^t, x_i^{t-1}, \dots, x_i^{t-k+1}) \in \mathbb{R}^{k \times 1}$ 和 $(x_i^t, x_i^{t-d}, \dots, x_i^{t-(k-1)d}) \in \mathbb{R}^{k \times 1}$ 表示进行需要卷积的向量（长度为 k ）； $W^g \in \mathbb{R}^{1 \times 1}$ ， $W^c \in \mathbb{R}^{k \times 1}$ ， $W_d^c \in \mathbb{R}^{k \times 1}$ 分别表示图卷积、因果卷积、因果膨胀卷积中需要学习的参数； k 表示卷积核的大小。当有多个卷积核时，卷积网络的输出维度将发生改变，例如， p 个卷积核可以将 $x_i \in \mathbb{R}^{n \times 1}$ 的维度变换为 $n \times p$ 。

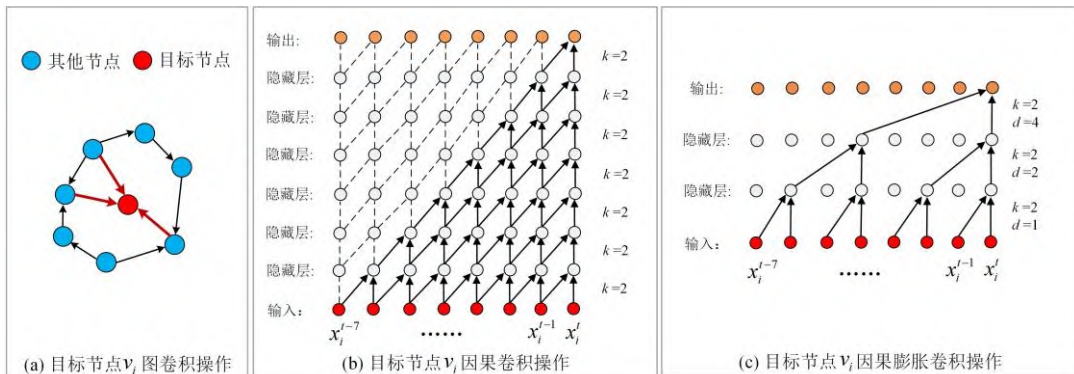


图 4 图卷积、因果卷积和因果膨胀卷积示意图

Fig. 4 Illustration of graph convolution, causal convolution and causal dilation convolution

在图卷积和因果膨胀操作的基础上, 本文进一步定义了 Causal-GCNM 模型的时空依赖模式挖掘层。时空依赖模式挖掘层的核心思想在于采用因果膨胀卷积操作和图卷积操作同时挖掘时空图序列的时空相关性。考虑到不同节点在不同时间窗口下具有不同的观测数据, 图卷积操作主要针对特定时间窗口进行, 例如 $\star^g v_i^t$ 表示对节点 v_i 在时间窗口 τ_t 中执行图卷积操作。具体而言, 集成图卷积和因果膨胀卷积操作的计算方法为

$$\left\{ \begin{aligned} (\star_d^c [x_i | x_{i:s}]) (t) &= \text{sum} \left(W_d^c \odot \begin{pmatrix} x_i^t & x_i^{t-d} & \dots & x_i^{t-(k-1)d} \\ x_{i:s}^t & x_{i:s}^{t-d} & \dots & x_{i:s}^{t-(k-1)d} \end{pmatrix} \right) \\ x_{i:s} &= \{x_{i:s}^1, \dots, x_{i:s}^{t-1}, x_{i:s}^t\} = \{\star^g v_i^1, \dots, \star^g v_i^{t-1}, \star^g v_i^t\} \\ \star^g v_i^t &= \sum_{j \in N_i} A_{ij} x_j^t W^g \end{aligned} \right. \quad (7)$$

式中 $x_i \in \mathbb{R}^{t \times 1}$ 表示节点 v_i 在时间维度上的观测值; $x_{i:s} \in \mathbb{R}^{t \times 1}$ 表示节点 v_i 在不同时间窗口中的图卷积结果, 即节点 v_i 附带的空间信息; A_{ij} 图中节点 v_i 和 v_j 的连接关系; $W_d^c \in \mathbb{R}^{k \times 2}$, $W^g \in \mathbb{R}^{1 \times 1}$ 分别表示因果膨胀卷积、图卷积中需要学习的参数。图 5 进一步展示了时空依赖模式挖掘层的详细流程。时间图序列首先经过图卷积神经网络获得包含邻域空间关系的时间图序列。然后, 两种时间图序列 (时间+空间) 经拼接后进入 N 层因果卷积网络获得最终的模型输出。参考文献[36]的思想, 每一层的因果卷积网络采用两个相同的因果膨胀卷积操作组成。此外, 与存在的时空图神经网络不同[7,18], 本文采用拼接的方式 (并联) 集成图卷积结果, 而非串联的方式集成图卷积结果。并联集成图卷积结果的优势在于可以在不割裂原有时间模式的前提下引入了额外的空间信息。

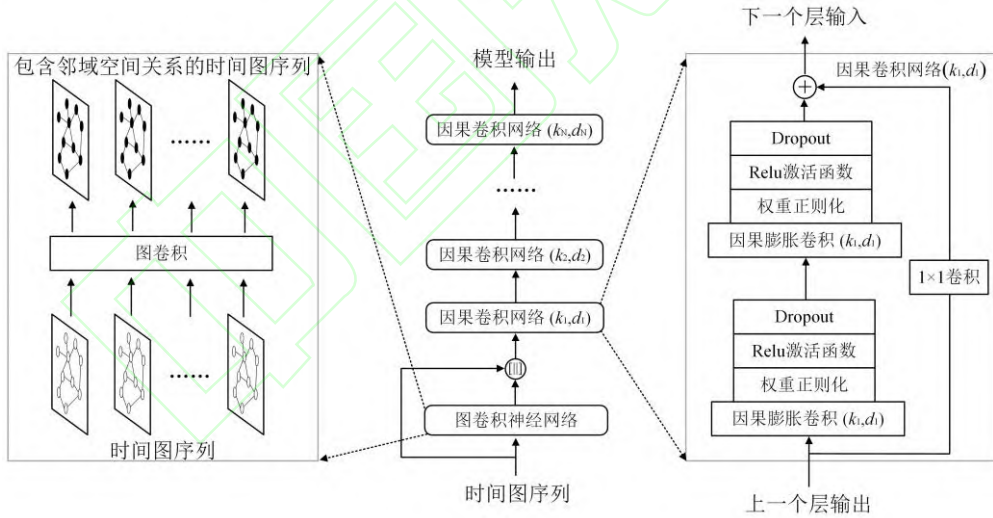


图 5 Causal-GCNM 模型的时空模式挖掘层示意图: 时空依赖模式挖掘层由单层图卷积神经网络 N 层因果卷积神经网络组成

Fig. 5 Illustration of the spatiotemporal dependent pattern mining layers of the Causal-GCNM model: The spatiotemporal dependent pattern mining layer consists of a single-layer graph convolutional network and N -layer causal convolutional network

2.3 模型优化

Causal-GCNM 模型是一种端到端的深度学习模型, 仅需要优化预测值与真值之间的均方误差即可获得最终的预测模型。在模型优化过程中, 为了减小模型的计算复杂度, 本文仅通过前 l 个时间窗口的时空数据 $\{x^j\}_{j=t-l+1}^t$ 预测将来的时空数据 \hat{x}^{t+1} 。此外, 考虑到数据缺失对预测结果的影响, 本文并不采用预测时间窗口 (时间窗口 τ_{t+1}) 的所有值优化 Causal-GCNM 模型的参数, 而是仅采用可观测数据优化 Causal-GCNM 模型的参数。具体而言, Causal-GCNM 模型采用的损失函数为

$$\mathcal{L}(\mathbf{W}) = \min_{\mathbf{W}} (\mathbf{x}^{t+1} \odot \mathbf{m}^{t+1} - \hat{\mathbf{x}}^{t+1} \odot \mathbf{m}^{t+1})^2 \quad (8)$$

式中： $\mathbf{x}^{t+1} \in \mathcal{R}^{n \times 1}$ 表示时间窗口 τ_{t+1} 内的时空数据观测值； $\hat{\mathbf{x}}^{t+1} \in \mathcal{R}^{n \times 1}$ 表示时间窗口 τ_{t+1} 内的时空数据预测值； $\mathbf{m}^{t+1} \in \mathcal{R}^{n \times 1}$ 用于标识 \mathbf{x}^{t+1} 中的观测值和缺失值； \odot 表示 Hadamard 积，即逐个元素相乘； \mathbf{W} 表示模型中优化的参数。当节点 v_i 在时间窗口 τ_{t+1} 内的数据缺失时， $x_i^{t+1} \odot m_i^{t+1}$ 和 $\hat{x}_i^{t+1} \odot m_i^{t+1}$ 同时为 0，即损失函数不累计节点 v_i 在时间窗口 τ_{t+1} 内的预测误差，从而达到了仅采用可观测数据优化 Causal-GCNM 模型参数的目的。

3 试验与讨论

3.1 数据获取

3.1.1 试验数据

我们采用三种时空数据集评估 Causal-GCNM 模型的性能，即交通流量数据、PM_{2.5} 监测数据和气温监测数据。

交通流量数据通过自动车辆识别（Automatic Vehicle Identification, AVI）技术获得，即通过摄像头拍摄的照片自动识别车辆的空间坐标。如图 6 (a)所示，本文共选取 67 个试验摄像头，以 5 分钟为间隔统计了单个摄像头的交通流量。

PM_{2.5} 监测数据来源于北京市 36 个空气质量监测站点^[40]，采样间隔为 60 分钟，图 6 (b)展示了监测站点的空间分布。

气温监测数据来源于哥白尼气候数据库^[41]，气温监测数据记录了内陆水域表面上方 2 米处的空气温度，采样间隔为 60 分钟。如图 6 (c)所示，本文共选取武汉及其周边地区的 64 个格网用于试验（格网分辨率为 $0.25^\circ \times 0.25^\circ$ ）。

表一为交通流量数据、PM_{2.5} 监测数据和气温监测数据的统计特征。交通流量数据和 PM_{2.5} 数据均存在不同程度的自然缺失。气温数据是经欧洲中期天气预报中心二次处理的数据，不存在自然缺失。

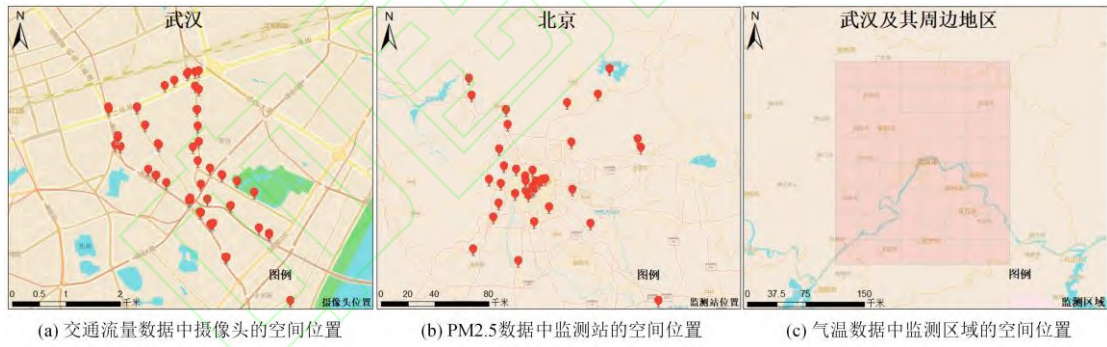


图 6 研究区域

Fig. 6 Study area

表 1 试验数据集说明

Tab. 1 Description of the datasets

数据集	交通流量数据	PM _{2.5} 数据	气温数据
所处位置	武汉	北京	武汉及其周边地区
时间窗口大小	5 分钟	60 分钟	60 分钟
计量单位	辆	微克每立方米	摄氏度
空间对象个数	67	36	64
时间窗口个数	8064	2952	2208
自然缺失率	8.94%	10.35%	0%
时间跨度	2021/3/1–2021/3/28	2014/5/1–2014/8/31	2018/6/1–2018/8/31

3.1.2 数据预处理

为支撑本文的研究，本文进一步针对时空数据进行了数据预处理，预处理过程主要如下：

(1) 交通流量数据、PM_{2.5} 监测数据、气温监测数据存在不同情况的自然缺失，自然缺失值会影响后续试验环境，即试验数据的缺失率。因此，本文首先补全了数据集中的自然缺失值。填补自然缺失值的过程仅仅是为了保证后续试验环境的一致，为了保证填补结果的精度，我们采用性能较优的 BiSTGN 模型^[42]补全了数据集中的自然缺失值。

(2) 借鉴于 Cui 等人^[33]的工作，基于两种缺失类型（随机缺失、块状缺失），分别以 20% 和 40% 为缺失率人工删除了部分时空数据，即人为将数据设置为 0。图 7 展示了三天的交通流量数据的缺失分布情况。

(3) 将人工处理的数据划分为训练数据集及测试数据集。依据二八准则，训练数据集占比 80%，测试数据集占比 20%。

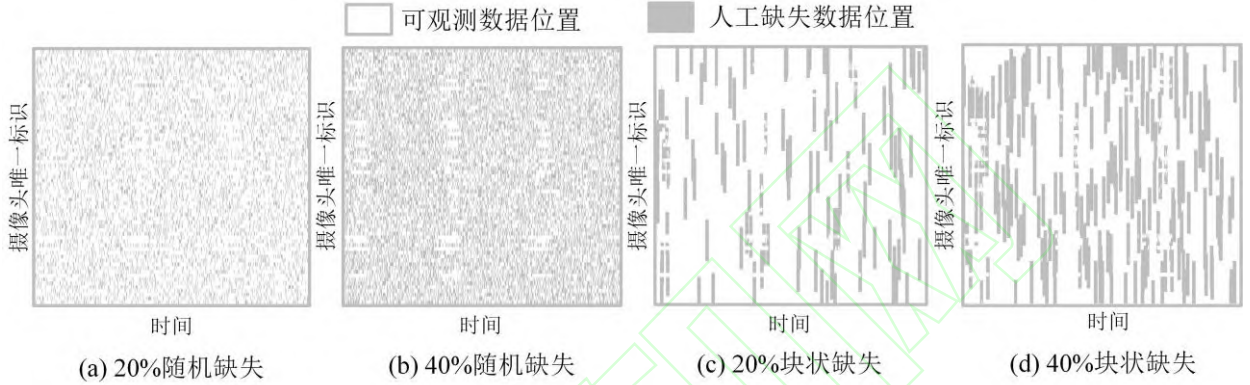


图 7 交通流量数据的人工缺失位置分布

Fig. 7 Missing location distribution of traffic volume data

3.2 定量评价指标

在时空预测中，一个关键问题是如何评估预测模型的性能。在本文中，平均绝对误差（Mean Absolute Error, MAE）、均方根误差（Root Mean Square Error, RMSE）作为定量指标来验证所提出模型的预测精度。其中 MAE 表示预测值与真实值之间的绝对误差的平均值，RMSE 表示预测值和真实值之间差异（残差）的样本标准偏差。MAE 和 RMSE 的计算方法分别为

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i^t - \hat{x}_i^t| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^t - \hat{x}_i^t)^2} \quad (10)$$

式中 x_i^t 表示节点 v_i 在未来第 t 个时间窗口内监测的真实值； \hat{x}_i^t 表示节点 v_i 在未来第 t 个时间窗口内的预测值； n 表示图中节点的总个数。

3.3 试验设置

Causal-GCNM 模型的超参数主要有 5 个：时间依赖步长、卷积核大小、因果卷积网络的层数、每层因果卷积网络中的卷积核个数、每层因果卷积网络中的膨胀因子。在模型训练过程中，本文主要使用控制变量法获得超参数的最优组合。表二展示了最终采用的超参数。

表 2 Causal-GCNM 模型的参数设置

Tab. 2 Parameter setting of Causal-GCNM model

参数名称	参数值
时间依赖步长	20
卷积核大小	2
因果卷积网络的层数	4
每层卷积核的个数	32, 64, 32, 1
每层膨胀因子	2, 4, 8, 16

3.4 与基准方法的对比分析

由于经典统计学模型在时空预测任务上的预测性能往往低于数据驱动模型的预测性能，本文主要将 Casual-GCNM 模型与流行的数据驱动方法进行了对比。本文采用的基线方法大致可以分为四类。第一类为不考虑缺失值的浅层机器学习模型（ST-KNN 模型^[13]）；第二类为不考虑缺失值的深度学习模型（GRU 模型^[43]、LSTM 模型^[44]、T-GCN 模型^[18]、ST-GCN 模型^[19]）；第三类为考虑缺失值的浅层机器模型（TRMF 模型^[32]、BTMF 模型^[31]）；第四类为考虑缺失值的深度学习模型（GRU-D 模型^[34]、LSTM-M 模型^[35]、SGMN 模型^[33]）。

3.4.1 非缺失数据集上的预测精度对比

表三首先展示了 Causal-GCNM 模型和基准方法在非缺失场景下的预测结果。整体而言，在非缺失场景下，四类模型的预测精度差异比较小，深度学习模型的预测性能略高于浅层机器学习的预测性能。预测模型在气温数据集上的预测精度高于预测模型在交通流数据集上的预测精度，预测模型在交通流数据集上的预测精度高于预测模型在 PM_{2.5} 数据集上的预测精度。主要原因是气温监测值在空间和时间相对比较平稳，使得预测模型更容易预测气温的走势。此外，T-GCN、ST-GCN 和 SGMN 等时空图预测模型的预测性能低于 TCN、GRU-D 和 LSTM-M 等时间预测模型的预测性能。造成这种结果的原因可能是 T-GCN、ST-GCN 和 SGMN 模型通过串联的方式集成了图卷积结果，从而割裂了原有的时间模式。与基准实验相比，Causal-GCNM 模型采用并联的方式集成图卷积操作，并没有割裂原有的时间模式。因此，Causal-GCNM 模型在非缺失场景下获得了最高的预测性能。

表 3 Causal-GCNM 与基准实验在无缺失场景下的比较结果（MAE\RMSE）

Tab. 3 Comparison results (in MAE/RMSE) of prediction performance between Causal-GCNM and baselines under non-missing scenario

模型	交通流数据	PM _{2.5} 数据	气温数据
ST-KNN	5.31/9.34	12.84/19.13	0.79/1.09
GRU	4.41/8.77	8.36/13.13	0.61/0.87
LSTM	4.21/7.73	8.61/13.63	0.61/0.86
T-GCN	6.95/11.86	14.91/21.31	0.80/1.06
ST-GCN	5.64/10.18	10.45/15.45	2.38/2.82
TRMF	4.79/8.02	11.26/15.98	0.73/0.98
BTMF	4.86/8.03	10.74/15.66	0.87/1.11
GRU-D	4.25/8.05	8.84/14.10	0.57/0.76
LSTM-M	4.24/8.15	8.85/14.08	0.59/0.82
SGMN	4.52/7.30	8.58/13.09	0.61/1.01
Causal-GCNM	4.17/6.84	8.25/12.87	0.51/0.69

3.4.2 缺失数据集上的预测精度对比

本文进一步比较了 Causal-GCNM 模型和基准方法在缺失场景下的预测结果，结果如表四和表五所示。在缺失场景下，四类预测模型表现出了明显的差异。其中考虑缺失值的预测模型要明显优于不考虑缺失值的预测模型。此外，在不同的缺失场景下，TRMF 和 BTMF 模型的预测性能比较稳定，而 GRU-D、LSTM-M、SGMN 模型的预测性能受缺失场景的影响较大。具体而言，GRU-D、LSTM-M、SGMN 模型在块

状缺失场景下的预测性能较差。主要原因为 GRU-D、LSTM-M 模型为纯时间序列模型，忽略了空间信息对缺失值的影响。虽然 TRMF 和 BTMF 模型的预测性能比较稳定，但在随机缺失场景下，其预测精度会略低于 GRU-D、LSTM-M 模型。与基准试验相比，Causal-GCNM 模型具有最高的预测精度。Causal-GCNM 模型获得高精度的原因主要有两点。首先，缺失数据处理组件使得 Causal-GCNM 模型具备了处理缺失值的能力。其次，图卷积操作使得 Causal-GCNM 模型可以在时间和空间两个维度挖掘数据集中的时空模式。总体而言，相较基准实验，Causal-GCNM 模型在具有明显的优势。

表 4 Causal-GCNM 与基准实验在随机缺失场景下的比较结果 (MAE\RMSE)

Tab. 4 Comparison results (in MAE/RMSE) of prediction performance between Causal-GCNM and baselines under random missing scenario

模型	缺失率: 20%			缺失率: 40%		
	交通流数据	PM _{2.5} 数据	气温数据	交通流数据	PM _{2.5} 数据	气温数据
ST-KNN	10.50/20.69	21.32/31.33	5.85/8.49	16.48/31.09	30.99/43.38	11.19/13.65
GRU	10.02/18.33	18.21/25.62	7.12/8.08	12.35/21.65	22.71/30.83	9.42/9.92
LSTM	10.17/18.78	18.72/26.29	7.33/8.23	12.58/22.12	23.19/39.14	9.57/10.06
T-GCN	12.21/21.55	19.92/28.17	6.95/7.16	16.26/27.45	27.45/37.28	10.75/11.95
ST-GCN	9.14/16.17	17.26/27.72	4.75/5.10	9.27/16.07	20.23/31.40	4.89/5.32
TRMF	5.06/9.44	13.02/19.29	0.85/1.19	5.74/14.22	15.47/24.71	0.87/1.19
BTMF	4.89/8.59	11.26/16.43	0.84/1.11	5.59/9.95	12.91/18.89	0.90/1.20
GRU-D	4.52/8.18	9.29/14.77	0.64/0.89	4.85/8.87	11.03/17.83	0.75/1.04
LSTM-M	4.51/8.89	9.39/14.81	0.67/0.93	4.79/9.43	11.14/17.87	0.81/1.10
SGMN	4.72/7.66	9.50/14.52	0.82/1.17	4.88/8.31	10.95/17.10	0.97/1.35
Causal-GCNM	4.44/7.32	9.59/14.42	0.62/0.79	4.63/7.80	10.53/15.97	0.63/0.81

表 5 Causal-GCNM 与基准实验在块状缺失场景下的比较结果 (MAE\RMSE)

Tab. 5 Comparison results (in MAE/RMSE) of prediction performance between Causal-GCNM and baselines under block missing scenario

模型	缺失率: 20%			缺失率: 40%		
	交通流数据	PM _{2.5} 数据	气温数据	交通流数据	PM _{2.5} 数据	气温数据
ST-KNN	16.52/32.30	24.93/39.59	7.29/11.75	20.95/39.55	26.11/39.31	10.61/14.64
GRU	10.11/24.95	18.59/35.45	6.27/11.94	14.35/33.36	22.79/40.25	9.97/15.43
LSTM	10.23/25.01	18.78/35.53	6.48/11.97	14.40/33.36	22.96/40.35	10.16/15.47
T-GCN	12.84/24.49	22.58/32.68	7.28/8.35	15.54/28.31	26.34/37.48	11.20/11.38
ST-GCN	11.45/23.07	20.77/37.44	5.86/7.31	13.89/28.73	25.06/42.71	7.44/9.62
TRMF	5.24/10.38	13.76/26.31	0.93/1.28	5.93/11.71	17.77/31.86	1.10/1.45
BTMF	5.15/9.26	12.53/18.14	0.89/1.15	5.82/10.61	14.19/22.67	0.96/1.28
GRU-D	5.29/10.95	11.22/18.66	0.85/1.29	6.40/13.89	14.13/21.32	1.08/1.61
LSTM-M	5.26/11.34	10.90/17.90	0.85/1.26	6.32/14.07	14.53/25.12	1.06/1.55
SGMN	6.66/15.00	13.22/22.82	1.94/3.57	8.57/20.59	16.60/27.63	3.00/4.90
Causal-GCNM	4.79/8.59	9.94/15.28	0.64/0.82	5.23/9.87	10.88/16.50	0.68/0.88

3. 4. 3 计算效率对比

在本小节，我们进一步分析了 Causal-GCNM 模型的计算效率。考虑到 Causal-GCNM 模型是一种深度学习模型，本文仅将 Causal-GCNM 模型与 GRU-D、LSTM-M 进行了对比，结果如表六所示。其中前向及反向传播时间主要用于反应 Causal-GCNM 模型离线训练速度的快慢，前向传播时间主要用于反应 Causal-GCNM 模型在线预测速度的快慢。由于单次运算可能存在一定的波动，本文利用 bootstrap 采样方法^[45]计

算了运算时间 95%的置信区间。与 GRU-D、LSTM-M 模型相比，Causal-GCNM 模型是一种非迭代类型的模型，其具备明显的运算优势。这表明 Causal-GCNM 模型也适用于实时性要求高的预测应用。此外，结果表明，在三种数据集上，三种模型的运算速度均是毫秒级别，原因是本文采用的数据集的规模比较小。在真实场景中，当数据集规模逐渐增大时，Causal-GCNM 模型的运算优势将会得到更加明显的体现。

表 6 Causal-GCNM 与基准实验计算效率在 40%块状缺失下的对比结果（MAE/RMSE）

Tab. 6 Comparison results (in MAE/RMSE) of calculation efficiency between Causal-GCNM and baselines under 40% block missing

模型		交通流数据 (毫秒)	PM _{2.5} 数据 (毫秒)	气温数据 (毫秒)
GRU-D	前向及反向传播	30.06±0.43	29.52±0.27	29.63±0.19
	前向传播	13.09±0.28	12.84±0.17	12.88±0.14
LSTM-M	前向及反向传播	31.22±1.05	31.18±0.49	30.70±0.36
	前向传播	13.15±0.20	13.17±0.28	12.97±0.15
Causal-GCNM	前向及反向传播	9.20±0.14	9.64±0.30	9.56±0.40
	前向传播	1.98±0.11	2.08±0.16	2.07±0.38

3.5 预测结果的定性分析

在本小节，折线图被用于定性的描述 Causal-GCNM 模型的预测性能。图 8 直观的展示了三种时空数据集在特定缺失率下的预测值和观测真值之间的差异。结果表明，在 40%块状缺失和 40%随机缺失场景下，观测真值与模型预测值较为接近，即 Causal-GCNM 模型准确的预测了时空数据的走势。这进一步证明了 Causal-GCNM 模型在缺失场景下依旧具备很好的预测性能。

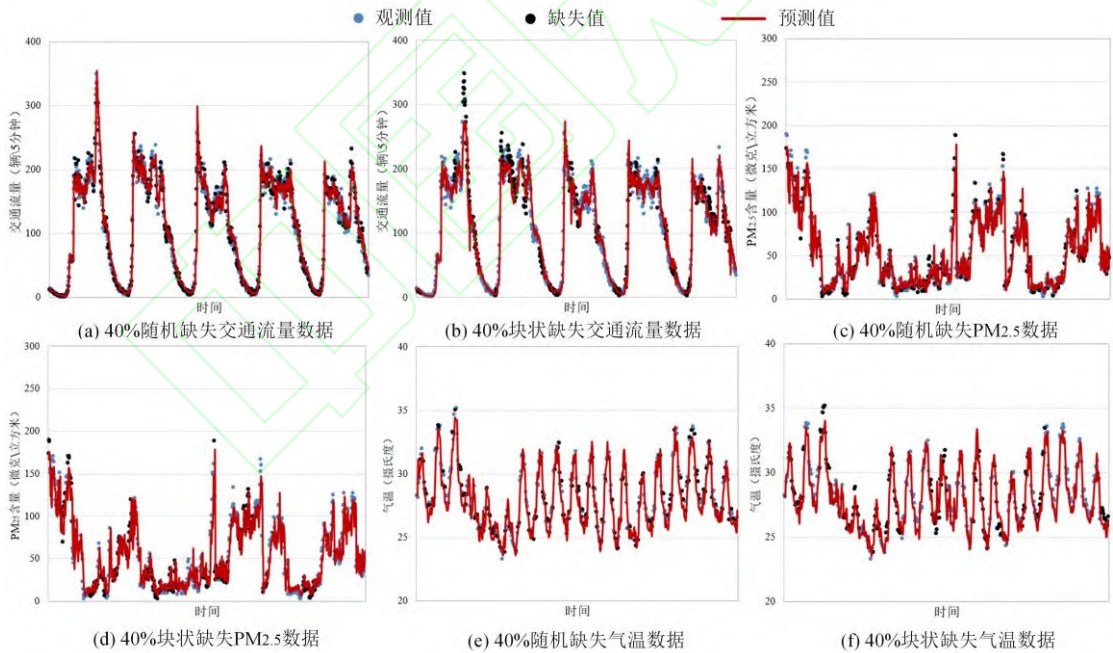


图 8 40%缺失率下预测值与真值之间的差异

Fig. 8 Difference between prediction value and actual value under 40% missing rate

3.6 不同组件对预测结果的影响

在本小节，本文进一步分析了不同组件对 Causal-GCNM 模型预测性能的影响，结果如表七所示。Causal-CN 模型表示经典的因果卷积网络，Causal-CNM 模型表示在 Causal-CN 基础上加入缺失数据处理的模型，Causal-GCNM 模型表示同时加入空间因素（图卷积）和缺失数据处理的模型，R 表示随机缺失数据集，B 表示块状缺失数据集。结果表明 Causal-CNM 模型的预测性能高于 Causal-CN 模型的预测性能，这表明引

入缺失处理组件的必要性。Causal-GCNM 的预测性能高于 Causal-CNM 模型的预测性能，这进一步证明了引入空间关系的必要性。此外，结果表明，在随机缺失模式下，Causal-GCNM 的预测性能受到缺失率的影响较小。原因是，在随机缺失模式下，缺失值往往不会连续缺失，即缺失值周围的数据依旧会存在。Causal-GCNM 模型可以有效的捕捉缺失值时空依赖关系从而获得精度较高的预测结果。在块状缺失模式下，Causal-GCNM 的预测性能受到缺失率的影响较大，Causal-GCNM 的预测性能会随着缺失率的上升逐渐下降。

表 7 不同组件对 Causal-GCNM 预测性能的影响 (MAE\RMSE)

Tab. 7 Impact of different components of Causal-GCNM on the prediction performance (MAE\RMSE)						
模型	缺失率: 20%			缺失率: 40%		
	交通流数据	PM _{2.5} 数据	气温数据	交通流数据	PM _{2.5} 数据	气温数据
Causal-CN (R)	9.41/17.80	20.20/28.27	7.01/8.08	13.34/24.88	24.44/33.54	9.47/10.12
Causal-CN (B)	10.83/24.11	19.56/35.58	6.73/11.88	14.54/31.74	24.21/41.07	10.32/15.36
Causal-CNM (R)	4.67/7.77	9.88/15.20	0.69/1.00	4.92/8.38	11.54/18.29	0.87/1.28
Causal-CNM (B)	5.43/10.85	11.36/18.40	0.96/1.55	6.49/14.10	14.80/25.10	1.26/1.94
Causal-GCNM (R)	4.44/7.32	9.59/14.42	0.62/0.79	4.63/7.80	10.53/15.97	0.63/0.81
Causal-GCNM (B)	4.79/8.59	9.94/15.28	0.64/0.82	5.23/9.87	10.88/16.50	0.68/0.88

4 结论与展望

时空预测是地理时空大数据挖掘的基础研究命题。目前存在的大多数时空预测模型仅在没有缺失值的时空数据上建模。然而，由于采集技术的各种问题和故障，无缺失值的情况在现实世界中很少存在。鉴于此，本文提出了一种新颖的考虑缺失值的 Causal-GCNM 模型用于时空数据预测任务。

在试验部分，本文使用三种真实的时空数据集（交通流数据集、PM_{2.5} 监测数据集、气温监测数据集）验证了 Causal-GCNM 模型的预测性能。首先，控制变量法被用于获得 Causal-GCNM 模型的最优参数组合。其次，本文将 Causal-GCNM 模型与存在的 10 种基线方法在三种数据集上进行了对比分析。实验结果表明，Causal-GCNM 模型在四种缺失条件（20%随机缺失、20%块状缺失、40%随机缺失、40%块状缺失）下仍然具有较好的预测性能，并在预测精度和计算效率两类指标上优于 10 种存在的基线方法。最后，我们分析了不同组件对 Causal-GCNM 预测性能的影响，进一步证明了 Causal-GCNM 模型适用于具有缺失值的时空预测。

本研究的尚有如下的局限性：（1）本文仅在低缺失率场景中验证了 Causal-GCNM 模型的预测性能，在实际场景中，高缺失率现象同样普遍存在；（2）本文仅验证了 Causal-GCNM 模型的单步预测性能，并未验证 Causal-GCNM 模型多步预测能力。针对上述问题，下一步工作重点将从两方面展开。首先，将 Causal-GCNM 模型应用到更高缺失率的场景。其次，进一步验证 Causal-GCNM 模型多步预测能力。此外，Causal-GCNM 模型是一种通用的顾及缺失值的时空预测模型，在将来有希望应用在更多的时空数据集中。

数据及代码声明：本文使用的数据及代码已经共享在 figshare 平台（<https://doi.org/10.6084/m9.figshare.17976803>）

参考文献:

[1] XU Fulong, Li Zishen, ZHANG Kefei, et al. An Investigation of Optimal Machine Learning Methods for the Prediction of ROTI[J]. Journal of Geodesy and Geoinformation Science, 2020, 3(2): 1–15. DOI:10.11947/j.JGGS.2020.0201.

[2] ZHANG Xiaohong, REN Xiaodong. Short-term Prediction of Ionospheric TEC Based on ARIMA Model[J]. Journal of Geodesy and Geoinformation Science, 2020, 2(1): 9–16. DOI:10.11947/j.JGGS.2019.0102.

[3] HAN Bing, ZHAO Guoze, WANG Lifeng, et al. Earthquake Electromagnetic Precursor Anomalies Detected by a

- New Ground-based Observation Network[J]. *Journal of Geodesy and Geoinformation Science*, 2021, 4(1): 116–123. DOI:10.11947/j.JGGS.2021.0114.
- [4] 王培晓, 张恒才, 王海波, 等. ST-CFSFDP:快速搜索密度峰值的时空聚类算法[J]. *测绘学报*, 2019, 48(11): 1380–1390. DOI:10.11947/j.AGCS.2019.20180538.
WANG Peixiao, ZHANG Hengcai, WANG Haibo, WU Sheng. Spatial-temporal clustering by fast search and find of density peaks[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(11): 1380-1390.
- [5] 赵琨, 罗力, 杨凤芸. 利用北斗 CORS 反演大气可降水量的精度分析[J]. *测绘科学*, 2021, 46(11): 12–17.
ZHAO Kun, LUO Li, YANG Fengyun. Accuracy analysis of precipitable water vapor retrieved by Beidou CORS[J]. *Science of Surveying and Mapping*, 2021, 46(11): 12–17.
- [6] 段功豪, 牛瑞卿, 赵艳南, 等. 基于动态指数平滑模型的降雨诱发滑坡预测[J]. *武汉大学学报(信息科学版)*, 2016, 41(07): 958–962.
DUAN Gonghao, NIU Ruiqing, ZHAO Yannan, et al. Rainfall-induced Landslide Prediction Based on Dynamic Exponential Smoothing Model [J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(07): 958–962.
- [7] 冯宁, 郭晟楠, 宋超, 等. 面向交通流量预测的多组件时空图卷积网络[J]. *软件学报*, 2019, 30(03): 759–769.
FENG Ning, GUO Shengnan, SONG Chao, et al. Multi-component Spatial-temporal Graph Convolution Networks for Traffic Flow Forecasting [J]. *Journal of Software*, 2019, 30(03): 759–769.
- [8] CHENG Shifen, LU Feng, PENG Peng. Short-Term Traffic Forecasting by Mining the Non-Stationarity of Spatiotemporal Patterns[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020: 1–19. DOI:10.1109/TITS.2020.2991781.
- [9] M. R. DEVRIES Phoebe, VIÉGAS Fernanda, WATTENBERG Martin, et al. Deep Learning of Aftershock Patterns Following Large Earthquakes[J]. *Nature*, 2018, 560(7720): 632–634. DOI:10.1038/s41586-018-0438-y.
- [10] 张显峰, 崔伟宏. 集成 GIS 和细胞自动机模型进行地理时空过程模拟与预测的新方法[J]. *测绘学报*, 2001(02): 148–155.
ZHANG Xianfeng, CUI Weihong. Integrating GIS with Cellular Automaton Model to Establish a New Approach for Spatio-temporal Process Simulation and Prediction[J]. *Acta Geodaetica et Cartographica Sinica*, 2001(02): 148–155.
- [11] 李静, 刘海砚, 郭文月, 等. 基于深度学习的人群活动流量时空预测模型[J]. *测绘学报*, 2021, 50(04): 522–531.
LI Jing, LIU Haiyan, GUO Wenyue, et al. A spatio-temporal network for human activity prediction based on deep learning[J]. *Acta Geodaetica et Cartographica Sinica*, 2021, 50(04): 522–531.
- [12] 吴华意, 黄蕊, 游兰, 等. 出租车轨迹数据挖掘进展[J]. *测绘学报*, 2019, 48(11): 1341–1356.
WU Huayi, HUANG Rui, YOU Lan, et al. A spatio-temporal network for human activity prediction based on deep learning[J]. *Acta Geodaetica et Cartographica Sinica*, 2019, 48(11): 1341–1356.
- [13] YU Bin, SONG Xiaolin, GUAN Feng, et al. K-Nearest Neighbor Model for Multiple-Time-Step Prediction of Short-Term Traffic Condition[J]. *Journal of Transportation Engineering*, 2016, 142(6): 04016018. DOI:10.1061/(ASCE)TE.1943-5436.0000816.
- [14] WU Shanhua, YANG Zhongzhen, ZHU Xiaocong, et al. Improved K-Nn for Short-Term Traffic Forecasting Using Temporal and Spatial Information[J]. *Journal of Transportation Engineering*, 2014, 140(7): 04014026. DOI:10.1061/(ASCE)TE.1943-5436.0000672.
- [15] ZHANG Junbo, ZHENG Yu, QI Dekang. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction[J], arxiv, <https://arxiv.org/abs/1610.00081>.
- [16] SHI Xingjian, GAO Zhihan, LAUSEN Leonard, et al. Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model[J], arxiv, <https://arxiv.org/abs/1706.03458>.
- [17] N.KIPF Thomas, WELLING Max. Semi-Supervised Classification with Graph Convolutional Networks[C/OL]/International Conference on Learning Representations. [2021–06–22].

<http://arxiv.org/abs/1609.02907>.

- [18] ZHAO Ling, SONG Yujiao, ZHANG Chao, et al. T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 21(9): 3848–3858. DOI:10.1109/TITS.2019.2935152.
- [19] YU Bing, YIN Haoteng, ZHU Zhanxing. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting[C/OL]//*Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*. Stockholm, Sweden: International Joint Conferences on Artificial Intelligence Organization, 2018: 3634–3640[2021–04–21]. <https://www.ijcai.org/proceedings/2018/505>. DOI:10.24963/ijcai.2018/505.
- [20] ZHANG Yang, CHENG Tao, REN Yibin, et al. A Novel Residual Graph Convolution Deep Learning Model for Short-Term Network-Based Traffic Forecasting[J]. *International Journal of Geographical Information Science*, 2020, 34(5): 969–995. DOI:10.1080/13658816.2019.1697879.
- [21] 方孟元, 唐炉亮. 基于低频 GNSS 轨迹的转向级城市交通信息精细预测[J]. *测绘学报*, 2021, 50(11): 1469. DOI:10.11947/j.AGCS.2021.20210252.
FANG Mengyuan, TANG Luliang. Fine-grained traffic information prediction at the turning-level based on low-frequency GNSS trajectory data[J]. *Acta Geodaetica et Cartographica Sinica*, 2021, 50(11): 1469.
- [22] 李明晓, 张恒才, 仇培元, 等. 一种基于模糊长短期神经网络的移动对象轨迹预测算法[J]. *测绘学报*, 2018, 47(12): 1660–1669.
LI Mingxiao, ZHANG Hengcai, QIU Peiyuan, et al. Predicting Future Locations with Deep Fuzzy-LSTM Network[J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(12): 1660–1669.
- [23] 樊子德, 龚健雅, 刘博, 等. 顾及时空异质性的缺失数据时空插值方法[J]. *测绘学报*, 2016, 45(04): 458–465.
FAN Zide, GONG Jianya, LIU Bo, et al. A Space-time Interpolation Method of Missing Data Based on Spatiotemporal Heterogeneity. *Acta Geodaetica et Cartographica Sinica*[J], 2016, 45(04): 458–465.
- [24] 樊子德. 异质时空数据插值方法研究[J]. *测绘学报*, 2017, 46(05): 668.
FAN Zide. Spatio-temporal Interpolation Methods for Heterogeneous Spatio-temporal Data[J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(05): 668.
- [25] 程诗奋, 彭澎, 张恒才, 等. 异质稀疏分布时空数据插值、重构与预测方法探讨[J]. *武汉大学学报(信息科学版)*, 2020, 45(12): 1919–1929.
CHENG Shifen, PENG Peng, ZHANG Hengcai, et al. Review of Interpolation, Reconstruction and Prediction Methods for Heterogeneous and Sparsely Distributed Geospatial Data[J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(12): 1919–1929.
- [26] YANG Jinming, PENG Zhongren, LIN Lei. Real-Time Spatiotemporal Prediction and Imputation of Traffic Status Based on LSTM and Graph Laplacian Regularized Matrix Factorization[J]. *Transportation Research Part C: Emerging Technologies*, 2021, 129: 103228. DOI:10.1016/j.trc.2021.103228.
- [27] CHENG Shifen, LU Feng, PENG Peng, et al. Short-Term Traffic Forecasting: An Adaptive ST-KNN Model That Considers Spatial Heterogeneity[J]. *Computers, Environment and Urban Systems*, 2018, 71: 186–198. DOI:10.1016/j.compenvurbsys.2018.05.009.
- [28] CHENG Shifen, LU Feng, PENG Peng, et al. Multi-Task and Multi-View Learning Based on Particle Swarm Optimization for Short-Term Traffic Forecasting[J]. *Knowledge-Based Systems*, 2019, 180: 116–132. DOI:10.1016/j.knosys.2019.05.023.
- [29] CAI Ling, JANOWICZ Krzysztof, MAI Gengchen, et al. Traffic Transformer: Capturing the Continuity and Periodicity of Time Series for Traffic Forecasting[J]. *Transactions in GIS*, 2020, 24(3): 736–755. DOI:10.1111/tgis.12644.
- [30] YI Zhiyan, LIU Xiaoyue Cathy, MARKOVIC Nikola, et al. Inferencing Hourly Traffic Volume Using Data-Driven Machine Learning and Graph Theory[J]. *Computers, Environment and Urban Systems*, 2021, 85: 101548. DOI:10.1016/j.compenvurbsys.2020.101548.
- [31] CHEN Xinyu, SUN Lijun. Bayesian Temporal Factorization for Multidimensional Time Series Prediction[J]. *IEEE*

- Transactions on Pattern Analysis and Machine Intelligence, 2021: 1–1. DOI:10.1109/TPAMI.2021.3066551.
- [32] YU H-F, RAO N, DHILLON I S. Temporal Regularized Matrix Factorization for High-Dimensional Time Series Prediction[C]//30th Conference on Neural Information Processing Systems (NIPS 2016). .
- [33] CUI Zhiyong, LIN Longfei, PU Ziyuan, et al. Graph Markov Network for Traffic Forecasting with Missing Data[J]. Transportation Research Part C: Emerging Technologies, 2020, 117: 102671. DOI:10.1016/j.trc.2020.102671.
- [34] CHE Zhengping, PURUSHOTHAM Sanjay, CHO Kyunghyun, et al. Recurrent Neural Networks for Multivariate Time Series with Missing Values[J]. Scientific Reports, 2018, 8(1): 6085. DOI:10.1038/s41598-018-24271-9.
- [35] TIAN Yan, ZHANG Kaili, LI Jianyuan, et al. LSTM-Based Traffic Flow Prediction with Missing Data[J]. Neurocomputing, 2018, 318: 297–305. DOI:10.1016/j.neucom.2018.08.067.
- [36] BAI Shaojie, KOLTER J.Zizo, KOLTUN Vladlen. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[J/OL]. arXiv:1803.01271 [cs], 2018[2021–05–05]. <http://arxiv.org/abs/1803.01271>.
- [37] YAN Jining, MU Lin, WANG Lizhe, et al. Temporal Convolutional Networks for the Advance Prediction of ENSO[J]. Scientific Reports, 2020, 10(1): 8055. DOI:10.1038/s41598-020-65070-5.
- [38] MA H, CHEN C, ZHU Q, et al. An ECG Signal Classification Method Based on Dilated Causal Convolution[J]. Computational and Mathematical Methods in Medicine, 2021, 2021: 1–10. DOI:10.1155/2021/6627939.
- [39] AYODEJI A, WANG Z, WANG W, et al. Causal Augmented ConvNet: A Temporal Memory Dilated Convolution Model for Long-Sequence Time Series Prediction[J/OL]. ISA Transactions, 2021[2022–01–27]. <https://www.sciencedirect.com/science/article/pii/S0019057821002810>. DOI:10.1016/j.isatra.2021.05.026.
- [40] ZHENG Y, YI X, LI M, et al. Forecasting Fine-Grained Air Quality Based on Big Data[C/OL]//Proceedings of the 21th SIGKDD conference on Knowledge Discovery and Data Mining. Proceedings of the 21th SIGKDD conference on Knowledge Discovery and Data Mining version. KDD 2015, 2015. <https://www.microsoft.com/en-us/research/publication/forecasting-fine-grained-air-quality-based-on-big-data/>.
- [41] HERBACH H, BELL B, BERRISFORD P, et al. (2018): ERA5 hourly data on single levels from 1979 to present[C/OL]//Copernicus Climate Change Service (C3S) Climate Data Store (CDS). [2022–01–03]. <https://www.baidu.com/>. DOI:10.24381/cds.adbb2d47.
- [42] WANG Peixiao, ZHANG Tong, ZHENG Yueming, et al. A Multi-View Bidirectional Spatiotemporal Graph Network for Urban Traffic Flow Imputation[J]. International Journal of Geographical Information Science, 2022. DOI:10.1080/13658816.2022.2032081.
- [43] CHUNG Junyoung, GULCEHRE Caglar, CHO KyungHyun, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling[J/OL]. arXiv:1412.3555 [cs], 2014[2021–07–20]. <http://arxiv.org/abs/1412.3555>.
- [44] 王培晓, 王海波, 傅梦颖, 等. 室内用户语义位置预测研究[J]. 地球信息科学学报, 2018, 20(12): 1689–1698. WANG Peixiao, WANG Haibo, FU Mengying, et al. Research on Semantic Location Prediction of Indoor Users[J]. Journal of Geo-Information Science, 2018, 20(12): 1689–1698.
- [45] 沈乐君, 游志胜, 李晓峰. 自助重要性采样用于实时多目标视觉跟踪[J]. 自动化学报, 2012, 38(10): 1663–1670. SHENG Lejun, YOU Zhisheng, LI Xiaofeng, Real-time Visual Tracking of Multiple Targets Using Bootstrap Importance Sampling[J]. Acta Automatica Sinica, 2012, 38(10): 1663–1670.

第一作者简介：王培晓（1994—），男，博士研究生，研究方向为时空数据分析、时空预测、轨迹分析等。

First author: WANG Peixiao（1994—），male, Doctoral candidate, majors in spatiotemporal data analysis, spatiotemporal prediction, trajectory analysis.

E-mail: peixiaowang@whu.edu.cn

通信作者：张彤

Corresponding author: ZHANG Tong E-mail: zhangt@whu.edu.cn